DOCUMENT RESUME

ED 471 364 EA 032 140

AUTHOR	Ross, Steven M.; Wang, L. Weiping; Alberg, Marty; Sanders, William L.; Wright, S. Paul; Stringfield, Sam
TITLE	Fourth-Year Achievement Results on the Tennessee Value-Added Assessment System for Restructuring Schools in Memphis.
PUB DATE	2001-04-00
NOTE	33p.; Paper presented at the meeting of the American Educational Research Association (Seattle, WA, April, 2001).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	*Change Strategies; Comparative Analysis; *Educational Change; Elementary Education; *School Effectiveness; Summative Evaluation; Systems Approach
IDENTIFIERS	Memphis City Schools TN; Tennessee Value Added Assessment System

ABSTRACT

This paper analyzes data on elementary schools in Memphis, Tennessee, that had reorganized 4 years prior to the study. The Tennessee Value-Added Assessment System (TVAAS) was used to provide performance scores showing students' year-to-year gains, whereas data analyzed for this paper were derived from scores on the TerraNova, the state-mandated achievement test, for five subjects over a 5-year period for grades 4-5. The research questions in this report include the following: How do restructuring schools compare in student-achievement gains to nonrestructuring schools over a 4-year period? and do different restructuring designs differ in achievement outcomes across the time period studied? Results indicate that in years 2 and 4 of implementation, restructured schools significantly outperformed nonrestructured schools in mathematics. In language, mathematics, science, and social studies, the restructured schools demonstrated significantly greater gain from pre-reform to post-reform than did nonrestructuring schools. The results suggest that incremental positive gains occurred in association with reforms, though most of the schools still had a long way to go to achieve high levels of student learning. Early analysis indicates that success in improving achievement appeared less a result of uniform reforms than of different designs appearing to work at particular schools and with groups of schools in given years. (Contains 18 references) (RJM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 471 364

Fourth-Year Achievement Results on the Tennessee Value-Added Assessment System for Restructuring Schools in Memphis

Steven M. Ross
L. Weiping Wang
Marty Alberg
The University of Memphis

William L. Sanders and S. Paul Wright
SAS, Inc.

Sam Stringfield
Johns Hopkins University

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
S. Ross
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Paper presented at the
April, 2001 meeting of the
American Educational Research Association
Seattle, WA

EA 032 140

BEST COPY AVAILABLE

Fourth-Year Achievement Results on the Tennessee Value-Added Assessment System for Restructuring Schools in Memphis

Executive Summary

In earlier studies, we examined the progress after two years and three years of the 25 elementary schools in Memphis City Schools (MCS) that began restructuring in 1995 compared to non-restructuring (control) schools in the district (see Ross, Sanders, Wright, & Stringfield, 1998; Ross, Wang, Sanders, Wright, & Stringfield, 1999). We also analyzed second-year data from 12 schools that began restructuring in 1996 (R96 cohort), the year following the initial wave. In the present study, we extended this prior research by analyzing fourth-year (1999) data from original restructuring elementary schools (R95), third-year data from the R96 schools, and second-year data from 19 R97 schools.

Achievement Analysis

The Tennessee Value-Added Assessment System (TVAAS) was employed to provide performance scores showing student year-to-year gains. Data analyzed in the present study were derived from scores on the TerraNova (a form of the CTBS-5), the state-mandated achievement test, for five subjects (math, reading, language, science, and social studies) over a five-year period for Grades 4-5. The index of student achievement, used in the present analyses is the *Cumulative % of Norm* mean (CPN). This statistic indicates across all grades reported the *percen. of the national (expected) gain.* For example, If School A had a CPN gain of 100% in math, it would have achieved at the national or expected level of achievement gain for that subject for that year.

Research Questions

The research questions addressed in this report are as follows:

1. How do restructuring schools compare in student achievement gains to non-restructuring schools over a four-year period?
2. Are achievement patterns similar for three cohorts of restructuring schools?
3. Do different restructuring designs differ in achievement outcomes across the time period studied?

Results and Conclusions

- In Year 2 (1997) and Year 4 (1999) of implementation, R95 schools significantly outperformed non-restructuring (NR) schools in Mathematics. In addition, the R95 schools showed nonsignificant advantages over the NR schools in all subjects averaged in all three post-reform years—1997, 1998, and 1999. Nonsignificant advantages in all subjects averaged were demonstrated by R96 schools in the second year (1998) and third year of reform (1999) and by R97 schools in their second year (1999).
- In Language, Mathematics, Science, Social Studies, and all subjects averaged, R95 schools demonstrated significantly greater gain from pre-reform (1995) to post-reform (1997/1998/1999) than did non-restructuring schools. Their mean advantage of 26.43 CPN points in all subjects averaged was associated with a fairly strong Effect Size of +0.79. R96 schools showed a moderate but nonsignificant advantage (*ES* = +0.58) over the NR schools in change scores from 1995/1996 to 1998/99. R97 schools showed a weak, nonsignificant advantage (*ES* = +0.33) over NR in change scores from 1995/1996/1997 to 1999.

- All three cohorts combined showed a moderately high, statistically significant advantage on all subjects averaged (*ES* = +0.67) over NR schools in pre- to post-reform change scores. The combined cohorts also significantly surpassed NR in Language (*ES* = +0.66), Mathematics (*ES* = +0.74), and Science (*ES* = +0.62).
- Across all cohorts and years of data, the following design cohorts demonstrated the most noticeable achievement gains: Accelerated Schools-95, Co-NECT-95, Roots & Wings-95, and Paidea-97.

Although MCS remains a low-performing district in comparison with state and national norms, it is also the highest-poverty and most urban district in Tennessee, with over 75% of its students qualifying for free or reduced-price lunch and over 80% being African American. Given the low achievement baseline, which has characterized MCS historically, it is unrealistic to expect that the district as a whole, as a result of any new curricula or programs, will attain the national median in achievement in one giant leap. If programs bring about positive changes, they will much more likely be incremental, reducing the gap from year to year. Thus, while the present results by no means indicate district success in achieving high levels of student learning, they do suggest that such incremental positive gains are occurring in association with the reforms.

With regard to design outcomes, there are too few schools represented in the present study to make confident comparisons or conclusions. In fact, success in improving achievement appears less a consistent product of certain designs than of different designs appearing to work effectively at particular schools and with groups of schools in given years. That is, for each of the eight designs examined, we can identify at least one school that experienced clear success in raising achievement and one that was not successful. Only

through continued research across other districts can the designs that prove most practical and beneficial be determined.

Fourth-Year Achievement Results on the Tennessee Value-Added Assessment System for Restructuring Schools in Memphis

Introduction

While there remains a lively debate among both scholars and practitioners as to the desirable locus of educational improvement, there is presently, through the Comprehensive School Reform Demonstration (CSRD) program ("Obey-Porter" legislation) and Title I School-Wide programs (Natriello & McDill, 1999), considerable impetus for implementing reforms through the implementation of whole-school change models. Given the extensiveness of current reform efforts nation-wide, it is surprising that there are very few reasonably rigorous studies attempting to discern the most practical alternatives for improving urban students' academic achievements (Nunnery, 1998). Beginning in the fall of 1995, Memphis City Schools (MCS) initiated one of America's most ambitious efforts to improve urban education by implementing eight diverse whole-school reform designs, including six New American Schools (NAS) designs, in 34 schools. A recent report by the RAND Corporation (Bodilly, Keltner, Purnell, Reichardt & Schulyer, 1998) indicated that Memphis was one of the two school districts in America that had made the greatest progress implementing NAS designs.

Despite the apparent progress in design implementation (Bodilly et al., 1998; Smith et al., 1998), the long-term question concerns whether these efforts are resulting in enhanced academic achievement for students. In two initial studies (Ross, Sanders, Wright, & Stringfield, 1998; Ross, Wang, Sanders, Wright, & Stringfield, 1999), we found positive

outcomes on state-mandated achievement tests for 25 restructuring elementary schools in Memphis following the second and third years of program implementation. The present study extended that research by analyzing two- to four-year longitudinal achievement outcomes for three cohorts of schools (total $n = 53$). It also reports summary results for individual designs.

Background For This Report

In 1997, we examined the progress of the 25 elementary schools in Memphis that began restructuring in 1995 (to be referred to as R95 schools). The team compared the reforming schools' gains on the Tennessee Comprehensive Assessment Program (TCAP) to those of 34 demographically-matched control or schools and the 40 remaining ("Other") non-restructuring elementary schools (see Ross et al., 1998). We then extended the study to include third-year (1998) data from those original 25 restructuring schools and second-year data from 12 schools that began restructuring in 1996 (R96 schools), the year following the initial wave (Ross et al., 1999). We restricted analyses to elementary schools due to inadequate numbers of middle schools and high schools in each cohort to allow for meaningful assessment of design-specific quantitative patterns.

In the latter study, we found that the advantages for restructuring schools over control or "non-restructuring" (NR) schools on the 1998 test were smaller than in 1997. The R95 schools, however, demonstrated significantly greater gain from pre-reform (1995) to post-reform (1997/1998) than did non-restructuring schools. In contrast, R96 schools showed only a small, nonsignificant advantage over the NR schools in change scores from 1995/1996 to 1998. The design showing the strongest and only significant effect across all subjects

averaged on the 1998 test was Co-NECT-95, while Roots &Wings-95 had a significant advantage over NR schools in Mathematics.

Analysis Procedures

The Tennessee Value-Added Assessment System (TVAAS) was employed to provide performance scores free of the biases normally associated with standardized test outcomes data (Sanders & Horn, 1995a, 1995b; Wright, Horn, & Sanders, 1997). TVAAS uses statistical mixed-model methodology to enable a multivariate, longitudinal analysis of student achievement, and to aggregate those data to the classroom and school levels. TVAAS scores have been demonstrated to yield estimates of student and teacher effects that are statistically independent of socioeconomic confoundings and do not require direct measures of these variables. Data analyzed in the present study were derived from scores on the TerraNova or CTBS-5 (CTB/MacMillan/McGraw Hill, 1997), the state-mandated achievement test, for five subjects (math, reading, language, science, and social studies) over a five-year period. In prior years, Tennessee schools employed TCAP, which was a form of the CTBS-4 (CTB/McGraw-Hill, 1990). An equating analysis conducted by Sanders and his staff in 1999 allowed for the conversion of TCAP scores to the TerraNova scale so that longitudinal effects could be determined (personal communication, William Sanders, April, 1999). In Memphis City Schools, CTBS/4 had been administered starting in Grade 3, thereby allowing for the computation of 1998 TVAAS scores in grades 4 and 5.

A particularly useful global index of student achievement, and the one emphasized in the present analyses, is the *Cumulative % of Norm* mean (CPN). This statistic indicates across all grades reported the *percent of the national (expected) gain attained* (Bratton, Horn, & Wright, 1996). For example, if School A had a CPN gain of 100% in fifth-grade math, it

would have achieved at the national or expected level of achievement gain for that subject for that grade in that year. If, however, its CPN in social studies was 80%, students' average gain in that subject would have been only .80 of the expected year-to-year gain. Using the TVAAS database of longitudinal CPN results by student and school, the present study examined the effects of the MCS restructuring initiative from spring, 1995 (pre-initiative) to spring, 1999. This was done by comparing the progress of restructuring elementary schools to non-restructuring schools in the same district.

The present R and NR samples changed from the Ross et al. (1999) study as a consequence of schools discontinuing designs or adopting new designs in the interim. Specifically, the R95 sample size decreased from 25 to 22 due to three schools discontinuing designs (one Co-NECT school and two Audrey Cohen College schools). The NR sample size decreased from 61 to 23 due to the 38 former NR schools implementing new designs. Mean percentile scores for 1995-1999 on all subjects averaged show that both R95 and NR schools in the present samples performed in the 32-43rd percentile range, clearly below the 50th percentile (the national average) and the State of Tennessee range from the 45-49th percentile, but typical of inner-city school districts serving large numbers of disadvantaged students (see Figure 1). The percentile scores indicate 7-9 point deficits for the R95 schools relative to NR schools from 1995 to 1998, and a 4 point difference in 1999. Thus, the R95 schools showed very modest progress in reducing their relative deficit in mean percentile score. However, the whole-school percentile mean, unlike the TVAAS scores, fails to take into account student mobility or to control for socioeconomic status or other variables that place certain schools at an advantage or disadvantage. For example, R schools had higher proportions of Title I schools than did the NR schools. Also, increases in the proportion of disadvantaged students

entering the school system each year has been associated with lower district percentile means at Grade 3 (the baseline year for computing TVAAS scores).

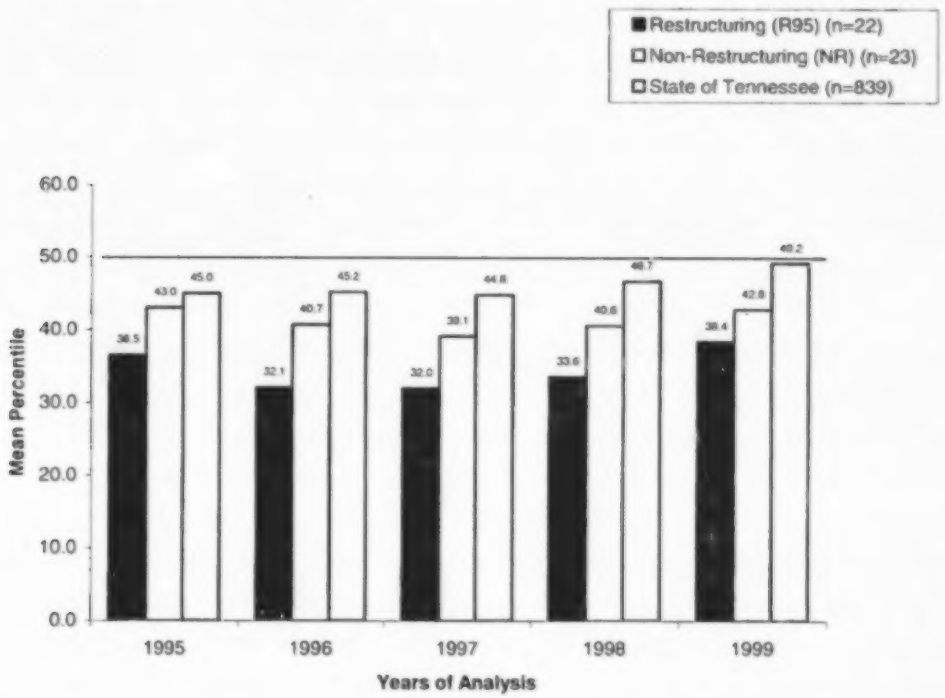

Figure 1. Memphis City Schools TVAAS Results (Mean Percentile) for R95 Schools vs. NR and State for 1995 – 1999.

BEST COPY AVAILABLE

Because of these factors, percentile outcomes may fail to reflect superior year-to year gains by individual students who remain at a given school. Finally, where whole-school *median* national percentiles (as frequently reported by the media) are used to judge schools' progress, an additional limitation is lack of sensitivity to student progress at higher or lower ends of the performance distribution. For example, if a school's median national percentile were 30, and many low-achieving students were moved from below the 15th percentile to close to the 25th (a clearly positive attainment), the median percentile would still remain at 30, erroneously suggesting zero growth. Thus, TVAAS scores are likely to provide a more valid measure of program effects.

Objectivity of the Study

In view of current federal and state initiatives promoting comprehensive school reform (CSR), competitions among design teams (see Herman, 1999) and criticisms by those who favor alternative approaches are likely to increase considerably. Therefore, it is important that studies of the effects of such designs be conducted and reported objectively. The present study arose out of an interest by the Center for Research in Educational Policy (CREP) researchers (Ross and associates), who have been evaluating school reform in that city and other districts for several years, and the University of Tennessee (UT) group (Sanders and associates, now at SAS), who developed the value-added system described above. Their mutual interests were in applying value-added assessment to perform a "cutting edge" evaluation of school reform effects. The study was conducted independently of Memphis City Schools, New American Schools, or any other organization. It was supported in part by an OERI grant awarded to CREP and by time donated by the two partner groups to

perform the research. All data analyses were conducted by the UT group using their statewide database.

Research Questions

The research questions addressed in this report are as follows:

1. How do restructuring schools compare in student achievement gains to non-restructuring schools over a four-year period?
2. Are achievement patterns similar for three cohorts of restructuring schools?
3. Do different restructuring designs differ in achievement outcomes across the time period studied?

Design

In the present study, we analyzed fourth-year (1999) data from 22 of the original 25 restructuring elementary schools (R95), third-year data from 12 schools that began restructuring in 1996 (R96), and second-year data from 19 schools that began restructuring in 1997 (R97). A total of 23 NR schools served as a comparison group. Table 1 summarizes the designs represented by cohort.

Table 1

Designs Represented in the Analysis by Year of Implementation

Design	1995	1996	1997	Total
Accel. Schools	2	3	5	10
ATLAS	2	1	0	3
Audrey Cohen College	2	0	0	2
Co-NECT	4	0	2	6
Exped. Learn/Out. Bound	2	2	0	4
Modern Red Schoolhouse	1	0	0	1
Paideia	1	2	2	5
Roots & Wings	8	4	10	22
TOTAL	**22**	**12**	**19**	**53**

Results

Descriptive Analyses

Figures 2 to 4 graphically depict the TerraNova-derived CPN means for all subjects averaged, from spring-1995 to spring-1999, for the R95, R96, and R97 cohorts, respectively. Each figure includes, for comparison purposes, comparable CPN means for NR schools and State of Tennessee schools (n = 839).

R95 results. As shown in Figure 2, in spring, 1995, prior to design implementation the following fall, the R95 schools (M = 106.2) were performing lower than NR schools (M = 121.5) and the state (M = 109.0). In 1996, after one year of implementation, R95 schools were still performing more poorly. However, in 1997, they showed advantages

relative to NR and the State of about 14 and 17 points, respectively. In 1998, the advantages were more modest, approximately 3 points in both comparisons. However, in the fourth year of restructuring (1999), the advantages increased to 16 and 19 points, respectively.

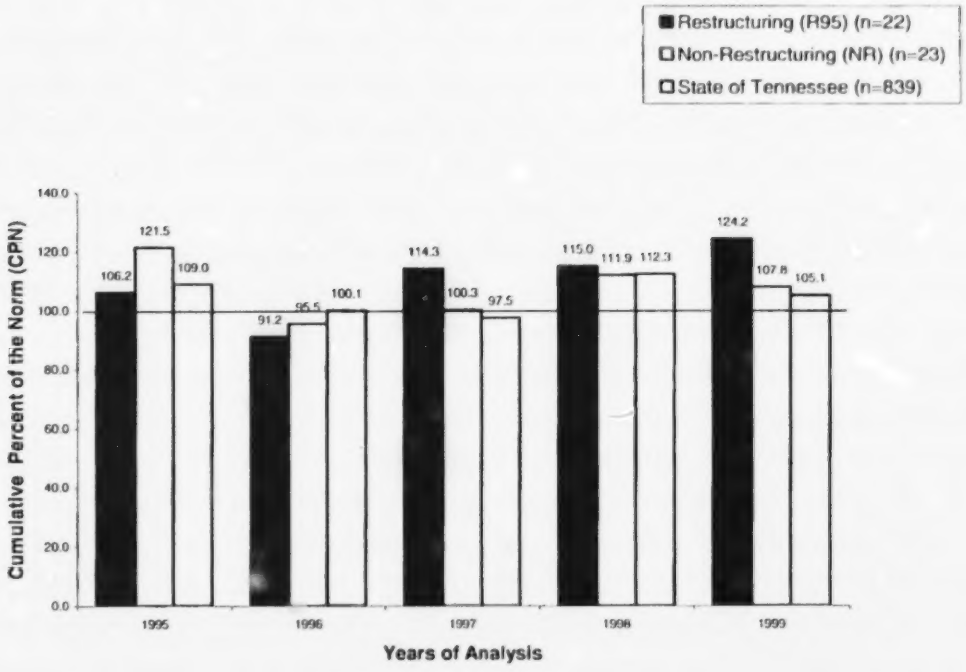

Figure 2. Memphis City Schools TVAAS Results 1995 Schools vs. NR and State for 1995 – 1999.

BEST COPY AVAILABLE

R96 results. As shown in Figure 3, the 12 R96 schools were performing lower than both NR schools (by about 9 CPN points) and State schools (by about 14 points) in 1996, the spring prior to implementation. In 1997, they were still performing lower, by 11 and 8 points, respectively. However, in 1998, after two years of implementation the R96 schools showed advantages of about 10 points over both NR and the State. In 1999, their advantage continued, although it decreased to 4 and 7 points, respectively.

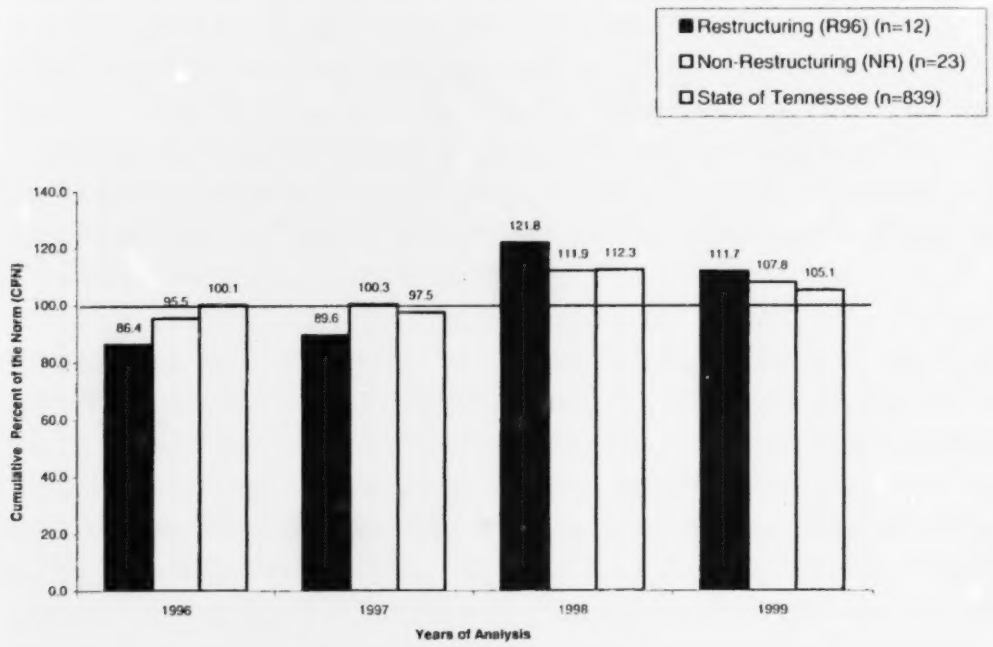

Figure 3. Memphis City Schools TVAAS Results for 1996 Schools vs. NR and State for 1996 – 1999.

R97 results. As shown in Figure 4, the 19 R97 schools performed comparably to NR and State schools in 1997, the year prior to restructuring. After one year (1998), they demonstrated fairly large deficits of about 16 points relative to both comparison groups. After two years (1999), they showed small to moderate advantages of advantages of 5 and 7 points, respectively.

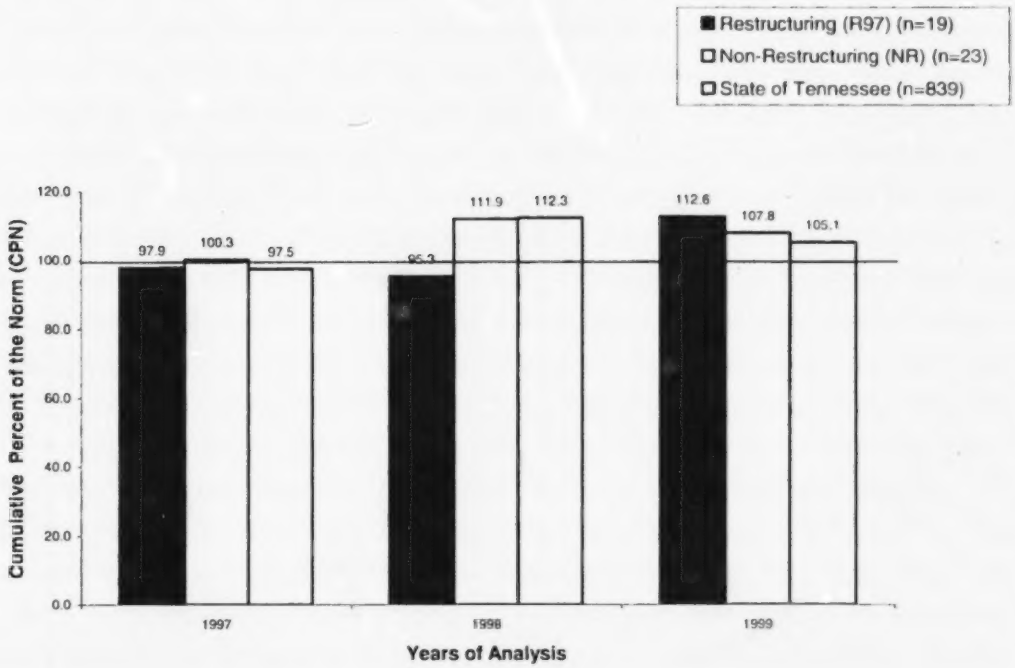

Figure 4. Memphis City Schools TVAAS Results for 1997 Schools vs. NR and State for 1997 – 1999.

BEST COPY AVAILABLE

Summary. Across the three cohorts of restructuring schools, the descriptive findings show a consistent pattern. In the year prior to beginning the reform, and especially in the first year of design implementation, the restructuring schools demonstrated lower levels of achievement gain than did NR and State schools. However, in the second year of implementation and beyond, the ordering was reversed, with the restructuring schools achieving relatively higher levels of gain.

The one post-reform year that all three cohorts have in common is 1999. To examine achievement outcomes in the aggregate for restructuring schools, we computed a weighted mean of 1999 CPNs on all subjects averaged for the R95, R96, and R97 cohorts. Figure 5 graphically depicts the outcome: the collective 53 R schools averaged 117.2 compared to means of 107.8 and 105.1 for the NR and State schools, respectively.

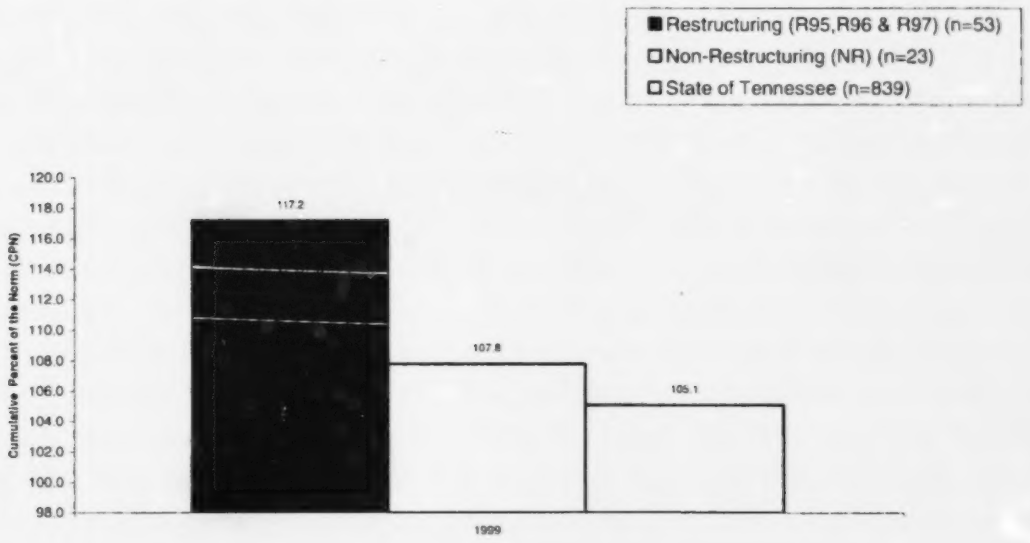

Figure 5. Cumulative (53 schools) Mean CPN vs. NR CPN for 1999.

Planned Inferential Comparisons

Inferential analyses examined differences between restructuring and non-restructuring schools overall and restructuring effects for individual designs. A probability level of .05 was used in determining significance. Because the unit of analysis is individual school, these comparisons generally have very low power and should therefore be considered highly conservative.

Year-by-year analyses. Comparisons between R and NR schools for each Cohort for each year yielded only a few significant differences:

- R95 schools scored significantly lower than NR schools in 1995, their pre-reform year, in Language ($p = .035$).
- R95 schools scored significantly higher than NR schools in Mathematics in 1997 ($p = .035$), their second year of reform, and in 1999 ($p = .008$).
- R97 schools scored significantly lower than NR schools in 1998, their first year of reform, in all subjects averaged ($p = .039$) and Reading ($p = .023$).

Pre- to post-reform change scores. To take pre-reform status and multiple year outcomes into account, we conducted planned comparisons of the differences between restructuring schools' and non-restructuring schools' CPN means prior to and after the reforms were implemented. These pre- to post-reform mean CPN differences will be called "change scores" throughout this report. For R95 schools, the pre-reform year was 1995, and the post-reform years were 1997, 1998 and 1999. For R96 schools, pre-reform was 1995 and 1996, and post-reform was 1998 and 1999. For R97 schools, pre-reform was 1995, 1996, and 1997, and post-reform was 1999.

Table 2 summarizes the results for each cohort on each subject and on all subjects averaged. The key statistic is the difference *estimate* reflecting the degree of pre- to post-reform change for the R95, R96, and R97 schools minus the same change mean for the NR schools. Thus, a positive estimate means that the R schools improved more (or declined less) than the NR schools over time; a negative estimate means the converse (i.e., NR schools improved more or declined less). The table also shows the significance of the *t* statistics comparing the R and NR change scores, and the effect size of the R advantage[1].

Table 2

Comparisons Between Restructuring and Non-restructuring Schools on Pre- to Post-Reform Change Means by Cohort and Subject

Subject	Cohort	Estimate	StdErr	t Value	ES
All	R95	26.43	10.08	2.62*	+0.79
	R96	15.20	9.40	1.62	+0.58
	R97	11.23	10.40	1.08	+0.33
	Combined	18.43	7.26	2.54*	+0.67
Reading	R95	10.61	11.83	0.90	+0.27
	R96	19.54	11.36	1.72	+0.61
	R97	4.31	14.35	0.30	+0.09
	Combined	10.37	9.13	1.14	+0.30

Table 2

Comparisons Between Restructuring and Non-restructuring Schools on Pre- to Post-Reform Change Means by Cohort and Subject (continued)

Subject	Cohort	Estimate	StdErr	t Value	ES
Language	R95	28.30	11.50	2.46*	+0.74
	R96	17.67	10.22	1.73	+0.62
	R97	10.05	10.35	0.97	+0.30
	Combined	19.35	7.79	2.48*	+0.66
Math	R95	28.97	8.76	3.31**	+1.00
	R96	14.43	8.73	1.65	+0.59
	R97	7.18	9.57	0.75	+0.23
	Combined	17.87	6.44	2.77**	+0.74
Science	R95	32.12	14.33	2.24*	+0.67
	R96	16.76	13.72	1.22	+0.44
	R97	21.53	15.56	1.38	+0.43
	Combined	24.84	10.46	2.37*	+0.62
Soc. Studies	R95	31.42	15.82	1.99*	+0.59
	R96	12.98	15.13	0.86	+0.31
	R97	12.98	16.22	0.80	+0.25
	Combined	20.63	11.68	1.77	+0.46

*$p < .05$ **$p < .01$

As shown in Table 2, R95 schools' pre- to post-reform change across all subjects averaged was 26.43 CPN points higher than the change for NR schools, a statistically significant difference ($p = .010$). Analyses of change scores for individual subjects, showed significant advantages for R95 over NR in Language ($M_{diff} = 28.3$, $p = .016$), Math ($M_{diff} = 28.97$, $p = .001$), Science ($M_{diff} = 32.12$, $p = .027$), and Social Studies ($M_{diff} = 31.42$, $p = .049$). None of the comparisons involving the R96 or R97 schools was significant, although all directionally favored the restructuring schools. Across all subjects averaged, the mean change score advantage was 15.20 for R96 schools and 11.23 for R97 schools.

Summary. Across all subjects averaged, R95 schools demonstrated greater gain from pre-reform (1995) to post-reform (1997/1998) than did non-restructuring schools. The mean advantage of 26.43 CPN points appears to be educationally important, especially when multiplied across 22 schools and the approximately 4,000 students served in total. Other statistically significant advantages were found for the R95 schools in Language, Math, Science, and Social Studies (but not in Reading). R96 and R97schools showed smaller, nonsignificant advantages over the non-restructuring schools in change scores from pre-reform to post reform years. The weighted change score averages for all restructuring schools combined were also computed and compared to the NR change means. As shown in Table 2, differences significantly favored the Combined restructuring schools in all subjects averaged and all individual subjects, except Reading.

Design Outcomes

To determine patterns for the individual designs implemented in Memphis, separate analyses were conducted comparing means for schools using particular designs to the NR school means. For a design to be included in this analysis, we required that it be employed in two or more MCS elementary schools. Table 3 summarizes the designs included by yearly cohort.

Table 3

Designs Represented by Two or More Schools by Cohort

Design	Cohort	Number of Schools
ATLAS	1995	2
Roots & Wings	1995	8
	1996	4
	1997	10
Co-NECT	1995	4
	1997	2
Accelerated Schools	1995	2
	1996	3
	1997	5
Expeditionary Learning	1995	2
	1996	2
Audrey Cohen College	1995	2
Paideia	1996	2
	1997	2

Because of the small numbers of schools representing each design, there is extremely low power for significance testing. Only Roots & Wings and Accelerated Schools have as many as 10 schools in total. Also, the fluctuation of results in any given year can give a misleading impression about the success of a design. We therefore decided to examine in

addition to separate year results for each design cohort, multiple-year change scores from pre-reform to post-reform (as just described for the overall R cohorts). For the change score differences, we also computed effect sizes.

In viewing the summary findings, readers should again consider that designs represented by larger numbers of schools (such as Roots and Wings) have much greater chance of showing significance where differences occur. Thus, we also looked at the absolute difference scores and include those comparisons where the differences appeared relatively sizable. Significant and/or sizeable differences occurred for the following individual year comparisons:

Accelerated Schools-95 < NR schools in 1996 (M_{diff} = -41.93, p = .045)

Co-NECT-95 > NR schools in 1998 (M_{diff} = +29.16, p = .115)

Roots and Wings-95 schools > NR schools in 1997 (M_{diff} = +37.67, p = .001)

Roots and Wings-95 schools > NR schools in 1999 (M_{diff} = +31.16, p = .016)

Listed below are pre- to post-reform mean change score differences between design schools and NR schools associated with effect sizes having an absolute value of .50 or greater on all subjects averaged:

Accelerated Schools-95 (M_{diff} = +38.78, p = .119, ES = +1.16)

Co-NECT-95 (M_{diff} = +26.86, p = .141, ES = +0.80)

ELOB-95 (M_{diff} = -40.95, p = .126, ES = -1.23)

Paideia-97 (M_{diff} = +60.63, p = .017, ES = +1.81)

Roots & Wings-95 (M_{diff} = +38.58, p = .007, ES = +1.15)

In a final series of analyses, we combined cohorts for designs having multiple cohorts of two or more schools. Two designs showed strong or significant effects:

Paideia-Combined > NR schools in all pre-reform to all post-reform years change (M_{diff} = +38.83, p = .018)

Roots and Wings-Combined > NR schools in all pre-reform to all post-reform years change (M_{diff} = +19.50, p = .033)

In summary, the most successful cohorts appear to have been Accelerated Schools-95, Co-NECT-95, Roots & Wings-95, and Paidea-97. ELOB-95 produced a negative change difference, not because of unusually low post-reform performance, but due to the 1995 pre-reform mean at one of the schools being extremely high. Given the small number of schools and year-to-year fluctuations in the means for these and other cohorts, the individual design results should be viewed cautiously.

Conclusions

Conclusions reached with respect to the main research questions are discussed in the following sections.

1. *How do restructuring schools compare in student achievement gains to non-restructuring schools over a four-year period?*

The results show that restructuring schools generally out-performed non-restructuring schools in year-to-year post-reform gain scores. The most striking results were for the 22 R95 schools in Years 2 (1997) and 4 (1999) of restructuring. Parallel but less dramatic (and nonsignificant) effects were found for the R96 and R97 cohorts. Specifically, R96 schools performed below NR schools prior to reform and after the first year of restructuring, but

above NR schools in post-reform years 2 and 3. R97 schools, similarly, showed advantages over NR in post-reform year 2.

Examinations of pre- to post-reform changes showed significantly greater *growth* for the R95 schools than for NR schools. Their mean advantage of +26.43 CPN points on all subjects averaged appears to represent an educationally important effect (*ES* = +0.79), especially when multiplied across the approximately 4,000 children served. Moderately strong and statistically significant advantages were found for the R95 schools in Language, Math, and Science, and Social Studies. R96 schools showed a small, nonsignificant advantage over the non-restructuring schools in change scores from 1995/1996 to 1998/1999, as did R97 schools in change scores from 1995/1996/1997 to 1999.

2. *Are achievement patterns similar for three cohorts of restructuring schools?*

In all of our analyses, the R95 schools have consistently shown the largest gains of the three restructuring cohorts. As we have speculated previously (Ross et al., 1999), the schools comprising R95 were the pioneers of the Memphis reform initiative and might have benefited from more committed and motivated staffs, stronger leadership, more attention or resources from the district and external design teams, and other factors. Over the period of this study (1995-99), all district schools were becoming increasingly exposed to strategies and interventions associated with comprehensive school reforms, such as standards-based learning and assessment and professional development in non-traditional (e.g., student-centered) teaching methods. Still, the NR schools have not gained as much over time as have the restructuring schools.

Although MCS remains a low-performing district in comparison with state and national norms, it is also the highest-poverty and most urban district in Tennessee, with over

75% of its students qualifying for free or reduced-price lunch and over 80% being African American. Given the low achievement baseline, which has characterized MCS historically, it is unrealistic to expect that the district as a whole, as a result of any new curricula or programs, will attain the national median in achievement in one giant leap. If programs bring about positive changes, they will much more likely be incremental, reducing the gap from year to year. Thus, while the present results by no means indicate district success in achieving high levels of student learning, they do suggest that such incremental positive gains are occurring in association with the reforms.

Identifying the critical influential elements of the reforms is beyond the scope of this study of student achievement. In other research, we have speculated that the designs have fostered teacher empowerment, professional development, cooperation and mutual planning, and exposure to additional resources (experts, processes, and materials) to support school governance, organization, curriculum, and instruction (see Ross, Henry et al., 1997; Smith et al., 1998). It is noteworthy, however, that all three cohorts showed the prototypic pattern that is usually associated with reform efforts, that is, little change or even deficits in achievement in the first year of implementation, but positive gains over time (see Levin, 1993). For all three of the cohorts, the "turnaround" period has been two years, which many educational reform experts would probably regard as unexpectedly brief given the demands and complexities of implementing comprehensive school reforms.

3. *Do restructuring designs differ in achievement outcomes?*

Given the relatively few schools using each of the designs and that implementation quality was not considered in the analyses, judgments about individual designs need to be made very cautiously. Yet, it seems important that where findings pertaining to particular

designs are reported, they be noted to allow for syntheses across studies to identify overall design effects in different contexts (Slavin, 1997). In the present study, success in improving achievement has appeared less a consistent product of certain designs than of different designs appearing to work effectively at particular schools and with groups of schools in given years. Specifically, the following for design cohorts demonstrated in our analyses the most noticeable achievement gains: Accelerated Schools-95, Co-NECT-95, Roots & Wings-95, and Paidea-97. Yet, we can find for each of the eight designs examined at least one school that experienced success in raising achievement and at least one that was not successful. While certain designs are likely to emerge over time as more practical and beneficial than others, in our research (e.g., Ross, Henry et al., 1997), the key factor for site-based success appears to be the fit of the design to the particular school and how the school uses the design as a framework and catalyst for improving its climate and educational programs.

References

Bodilly, S., Keltner, B., Purnell, S., Reichardt, R., & Schuyler, G. (March, 1998). *Lessons from New American Schools' scale-up phase: Prospects for bringing designs to multiple schools.* (MR-942.0-NAS). Santa Monica, CA: RAND.

Bratton, S. E., Jr., Horn, S. E., & Wright, S. P. (1996). *Using and interpreting Tennessee's value-added assessment system: A primer for teachers and principals.* Knoxville, TN: Value-Added Research and Assessment Center, University of Tennessee.

CTB/McGraw-Hill (1990). *Comprehensive test of basic skills (4th ed.).* Spring Norms Book. Monterey, CA: CTB/MacMillan/McGraw-Hill.

CTB/McGraw-Hill (1997). *TerraNova Technical Bulletin I.* Monterey, CA: CTB/McGraw-Hill.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Herman, R. (1999). *An educators' guide to schoolwide reform.* Arlington, VA: Educational Research Service.

Levin, H. (1993). *Learning from Accelerated Schools.* Unpublished paper. Palo Alto, CA: Stanford University.

Natriello, G., & McDill, E. L. (1999). Title I: From funding mechanism to educational reform. In G. Orfield & E. H. DeBray (Eds.), *Hard work for good schools: Facts not fads in Title I reform* (pp. 31-45). Cambridge, MA: The Civil Rights Project, Harvard University.

Nunnery, J. A. (1998). Reform ideology and the locus of development problem in educational restructuring: Enduring lessons from studies of educational innovation. *Education and Urban Society, 30*(3), 277-295.

Ross, S.M., Henry, D., Phillipsen, L., Evans, K., Smith, L., Buggey, T. (1997). Matching restructuring programs to schools: Selection, negotiation, and preparation. *School Effectiveness and School Improvement, 8*(1), 45-71.

Ross, S. M., Sanders, W. L., Wright, S. P., & Stringfield, S. (June, 1998). *The Memphis restructuring initiative: Achievement results for years 1 and 2 on the Tennessee Value-Added Assessment System (TVAAS).* Memphis, TN: Center for Research in Educational Policy, The University of Memphis.

Ross, S. M., Wang, W., Sanders, W. L., Wright, S. P., & Stringfield, S. (1999). *Two- and three-year achievement results on the Tennessee value-added assessment system for restructuring schools in Memphis.* Memphis, TN: Center for Research in Educational Policy: The University of Memphis.

Sanders, W. L., & Horn, S. P. (1995a). Educational assessment reassessed: The usefulness of standardized and alternative measures of student achievement as indicators for the assessment of educational outcomes. *Educational Policy Analysis Archives, 3*(6).

Sanders, W. L., & Horn, S. P. (1995b). The Tennessee Value-Added Assessment System (TVAAS): Mixed model methodology in educational assessment. In A. J. Shinkfield & D. Stufflebeam (Eds.), *Teacher evaluation: Guide to effective practice* (pp. 337-350). Boston: Kluwer.

Slavin, R. E. (1997). Design competitions: A proposal for a new federal role in educational research and development. *Educational Researcher*, 26(1), 22-28.

Smith, L., Ross, S., McNelis, M., Squires, M., Wasson, R., Maxwell, S., Weddle, K., Nath, L., Grehan, A., & Buggey, T. (1998). The Memphis restructuring initiative: Analyses of activities and outcomes that impact implementation success. *Education and Urban Society, 30*(3), 296-325.

Wright, S.P., Horn, S.P., & Sanders, W.L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*, 57-67.

Bodilly, S., Keltner, B., Purnell, S., Reichardt, R., & Schuyler, G. (March, 1998). *Lessons from New American Schools' scale-up phase: Prospects for bringing designs to multiple schools.* (MR-942.0-NAS). Santa Monica, CA: RAND.

Footnotes

[1]The effect size index computed is Cohen's *d* (Cohen, 1988, p. 20), defined as $(M_1 - M_2)/s$.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Fourth Year Achievement Results on the Tennessee Value-Added...	
Author(s): Ross, Sanders, et. al.	
Corporate Source:	Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1 ↑

[X]

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A ↑

[ ]

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B ↑

[ ]

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature:	Printed Name/Position/Title: Professor	
Organization/Address: Univ. of Memphis	Telephone: 901-678-3413	FAX: 901-678-4257
	E-Mail Address: smross@memphis.edu	Date: 6/7/02

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC Clearinghouse on Educational Management
~~1787 Agate Street~~
5207 University of Oregon
Eugene, OR 97403-5207
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: info@ericfac.piccard.csc.com
WWW: http://ericfacility.org

EFF-088 (Rev. 2/2000)